

## SEQUENCE LISTING

<110> Zavada, Jan  
Pastorekova, Silvia  
Pastorek, Jaromir

<120> MN Gene and Protein

<130> D-0021.5B-2

<140>  
<141>

<150> 09/178,115

<151> 1998-10-23

<160> 116

<170> PatentIn Ver. 2.1

<210> 1

<211> 1522

<212> DNA

<213> HUMAN

<220>

<221> CDS

<222> (13) .. (1389)

<220>

<221> mat\_peptide

<222> (124) .. (1389)

<400> 1

acagtcagcc gc atg gct ccc ctg tgc ccc agc ccc tgg ctc cct ctg ttg 51  
Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu  
-35 -30 -25

atc ccg gcc cct gct cca ggc ctc act gtg caa ctg ctg tca ctg 99  
Ile Pro Ala Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Ser Leu  
-20 -15 -10

ctg ctt ctg atg cct gtc cat ccc cag agg ttg ccc cggt atg cag gag 147  
Leu Leu Met Pro Val His Pro Gln Arg Leu Pro Arg Met Gln Glu  
-5 -1 1 5

gat tcc ccc ttg gga gga ggc tct tct ggg gaa gat gac cca ctg ggc			195
Asp Ser Pro Leu Gly Gly Ser Ser Gly Glu Asp Asp Pro Leu Gly			
10	15	20	
gag gag gat ctg ccc agt gaa gag gat tca ccc aga gag gag gat cca			243
Glu Glu Asp Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp Pro			
25	30	35	40
ccc gga gag gag gat cta cct gga gag gag gat cta cct gga gag gag			291
Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu			
45	50	55	
gat cta cct gaa gtt aag cct aaa tca gaa gaa gag ggc tcc ctg aag			339
Asp Leu Pro Glu Val Lys Pro Lys Ser Glu Glu Gly Ser Leu Lys			
60	65	70	
tta gag gat cta cct act gtt gag gct cct gga gat cct caa gaa ccc			387
Ieu Glu Asp Leu Pro Thr Val Glu Ala Pro Gly Asp Pro Gln Glu Pro			
75	80	85	
cag aat aat gcc cac agg gac aaa gaa ggg gat gac cag agt cat tgg			435
Gln Asn Asn Ala His Arg Asp Lys Glu Gly Asp Asp Gln Ser His Trp			
90	95	100	
ccg tat gga ggc gac ccg ccc tgg ccc cggt gtc tcc cca gcc tgc gcg			483
Arg Tyr Gly Gly Asp Pro Pro Trp Pro Arg Val Ser Pro Ala Cys Ala			
105	110	115	120
ggc cgc ttc cag tcc ccg gtg gat atc cgc ccc cag ctc gcc gcc ttc			531
Gly Arg Phe Gln Ser Pro Val Asp Ile Arg Pro Gln Leu Ala Ala Phe			
125	130	135	
tgc ccg gcc ctg cgc ccc ctg gaa ctc ctg ggc ttc cag ctc ccg ccg			579
Cys Pro Ala Leu Arg Pro Leu Glu Leu Gly Phe Gln Leu Pro Pro			
140	145	150	
ctc cca gaa ctg cgc ctg cgc aac aat ggc cac agt gtg caa ctg acc			627
Leu Pro Glu Leu Arg Leu Arg Asn Asn Gly His Ser Val Gln Leu Thr			
155	160	165	
ctg cct cct ggg cta gag atg gct ctg ggt ccc ggg cggt gag tac ccg			675
Leu Pro Pro Gly Leu Glu Met Ala Leu Gly Pro Gly Arg Glu Tyr Arg			
170	175	180	

gct ctg cag ctg cat ctg cac tgg ggg gct gca ggt cgt ccg ggc tcg Ala Leu Gln Leu His Leu His Trp Gly Ala Ala Gly Arg Pro Gly Ser 185	190	195	200	723
gag cac act gtg gaa ggc cac cgt ttc cct gcc gag atc cac gtg gtt Glu His Thr Val Glu Gly His Arg Phe Pro Ala Glu Ile His Val Val 205	210	215		771
cac ctc agc acc gcc ttt gcc aga gtt gac gag gcc ttg ggg cgc ccg His Leu Ser Thr Ala Phe Ala Arg Val Asp Glu Ala Leu Gly Arg Pro 220	225	230		819
gga ggc ctg gcc gtg ttg gcc ttt ctg gag gag ggc ccg gaa gaa Gly Gly Leu Ala Val Leu Ala Ala Phe Leu Glu Glu Gly Pro Glu Glu 235	240	245		867
aac agt gcc tat gag cag ttg ctg tct cgc ttg gaa gaa atc gct gag Asn Ser Ala Tyr Glu Gln Leu Leu Ser Arg Leu Glu Glu Ile Ala Glu 250	255	260		915
gaa ggc tca gag act cag gtc cca gga ctg gac ata tct gca ctc ctg Glu Gly Ser Glu Thr Gln Val Pro Gly Leu Asp Ile Ser Ala Leu Leu 265	270	275	280	963
ecc tct gac ttc agc cgc tac ttc caa tat gag ggg tct ctg act aca Pro Ser Asp Phe Ser Arg Tyr Phe Gln Tyr Glu Gly Ser Leu Thr Thr 285	290	295		1011
ccg ccc tgt gcc cag ggt gtc atc tgg act gtg ttt aac cag aca gtg Pro Pro Cys Ala Gln Gly Val Ile Trp Thr Val Phe Asn Gln Thr Val 300	305	310		1059
atg ctg agt gct aag cag ctc cac acc ctc tct gac acc ctg tgg gga Met Leu Ser Ala Lys Gln Leu His Thr Leu Ser Asp Thr Leu Trp Gly 315	320	325		1107
cct ggt gac tct cgg cta cag ctg aac ttc cga gcg acg cag cct ttg Pro Gly Asp Ser Arg Leu Gln Leu Asn Phe Arg Ala Thr Gln Pro Leu 330	335	340		1155
aat ggg cga gtg att gag gcc tcc ttc cct gct gga gtg gac agc agt Asn Gly Arg Val Ile Glu Ala Ser Phe Pro Ala Gly Val Asp Ser Ser 345	350	355	360	1203

cct	cg	g	c	g	t	g	a	g	c	ca	gt	c	a	g	at	tcc	tgc	ctg	g	c	t	g	g	tt	1251					
Pro	Arg	Ala	Ala	Glu	Pro	Val	Gln	Leu	Asn	Ser	Cys	Leu	Ala	Ala	Gly															
																365		370									375			
gac	atc	cta	gcc	ctg	gtt	ttt	ggc	ctc	ctt	ttt	gct	gtc	acc	agc	gtc											1299				
Asp	Ile	Leu	Ala	Leu	Val	Phe	Gly	Leu	Leu	Phe	Ala	Val	Thr	Ser	Val															
																380		385									390			
gcg	ttc	ctt	gtg	cag	atg	aga	agg	cag	cac	aga	agg	gga	acc	aaa	ggg											1347				
Ala	Phe	Leu	Val	Gln	Met	Arg	Arg	Gln	His	Arg	Arg	Gly	Thr	Lys	Gly															
																395		400									405			
ggt	gtg	agc	tac	cgc	cca	gca	gag	gta	gcc	gag	act	gga	gcc												1389					
Gly	Val	Ser	Tyr	Arg	Pro	Ala	Glu	Val	Ala	Glu	Thr	Gly	Ala																	
																410		415									420			
tagaggctgg	atcttgaga	atgtgagaag	ccagccagag	gcatctgagg	gggagccggt																					1449				
aactgtcctg	tcctgctcat	tatgccactt	ccttttaact	gccaaagaaat	tttttaaaat																					1509				
aaatatttat	aat																								1522					
<210>	2																													
<211>	459																													
<212>	PRT																													
<213>	HUMAN																													
<400>	2																													
Met	Ala	Pro	Leu	Cys	Pro	Ser	Pro	Trp	Leu	Pro	Leu	Leu	Ile	Pro	Ala															
																-35		-30									-25			
Pro	Ala	Pro	Gly	Leu	Thr	Val	Gln	Leu	Leu	Leu	Ser	Leu	Leu	Leu	Leu															
																-20		-15									-10			
Met	Pro	Val	His	Pro	Gln	Arg	Leu	Pro	Arg	Met	Gln	Glu	Asp	Ser	Pro															
																-5		-1	1								5		10	
Leu	Gly	Gly	Gly	Ser	Ser	Gly	Glu	Asp	Asp	Pro	Leu	Gly	Glu	Glu	Asp															
																15		20									25			
Leu	Pro	Ser	Glu	Glu	Asp	Ser	Pro	Arg	Glu	Glu	Asp	Pro	Pro	Gly	Glu															
																30		35									40			
Glu	Asp	Leu	Pro	Gly	Glu	Glu	Asp	Leu	Pro	Gly	Glu	Glu	Asp	Leu	Pro															
																45		50									55			

Glu Val Lys Pro Lys Ser Glu Glu Glu Gly Ser Leu Lys Leu Glu Asp  
60 65 70 75

Leu Pro Thr Val Glu Ala Pro Gly Asp Pro Gln Glu Pro Gln Asn Asn  
80 85 90

Ala His Arg Asp Lys Glu Gly Asp Asp Gln Ser His Trp Arg Tyr Gly  
95 100 105

Gly Asp Pro Pro Trp Pro Arg Val Ser Pro Ala Cys Ala Gly Arg Phe  
110 115 120

Gln Ser Pro Val Asp Ile Arg Pro Gln Leu Ala Ala Phe Cys Pro Ala  
125 130 135

Leu Arg Pro Leu Glu Leu Leu Gly Phe Gln Leu Pro Pro Leu Pro Glu  
140 145 150 155

Leu Arg Leu Arg Asn Asn Gly His Ser Val Gln Leu Thr Leu Pro Pro  
160 165 170

Gly Leu Glu Met Ala Leu Gly Pro Gly Arg Glu Tyr Arg Ala Leu Gln  
175 180 185

Leu His Leu His Trp Gly Ala Ala Gly Arg Pro Gly Ser Glu His Thr  
190 195 200

Val Glu Gly His Arg Phe Pro Ala Glu Ile His Val Val His Leu Ser  
205 210 215

Thr Ala Phe Ala Arg Val Asp Glu Ala Leu Gly Arg Pro Gly Gly Leu  
220 225 230 235

Ala Val Leu Ala Ala Phe Leu Glu Glu Gly Pro Glu Glu Asn Ser Ala  
240 245 250

Tyr Glu Gln Leu Leu Ser Arg Leu Glu Glu Ile Ala Glu Glu Gly Ser  
255 260 265

Glu Thr Gln Val Pro Gly Leu Asp Ile Ser Ala Leu Leu Pro Ser Asp  
270 275 280

Phe Ser Arg Tyr Phe Gln Tyr Glu Gly Ser Leu Thr Thr Pro Pro Cys  
285 290 295

Ala Gln Gly Val Ile Trp Thr Val Phe Asn Gln Thr Val Met Leu Ser  
300 305 310 315

Ala Lys Gln Leu His Thr Leu Ser Asp Thr Leu Trp Gly Pro Gly Asp  
320 325 330

Ser Arg Leu Gln Leu Asn Phe Arg Ala Thr Gln Pro Leu Asn Gly Arg  
335 340 345

Val Ile Glu Ala Ser Phe Pro Ala Gly Val Asp Ser Ser Pro Arg Ala  
350 355 360

Ala Glu Pro Val Gln Leu Asn Ser Cys Leu Ala Ala Gly Asp Ile Leu  
365 370 375

Ala Leu Val Phe Gly Leu Leu Phe Ala Val Thr Ser Val Ala Phe Leu  
380 385 390 395

Val Gln Met Arg Arg Gln His Arg Arg Gly Thr Lys Gly Gly Val Ser  
400 405 410

Tyr Arg Pro Ala Glu Val Ala Glu Thr Gly Ala  
415 420

<210> 3

<211> 29

<212> DNA

<213> HUMAN

<400> 3

cgccccagtgg gtcatcttcc ccagaagag

29

<210> 4

<211> 19

<212> DNA

<213> HUMAN

<400> 4

ggaatcctcc tgcatccgg

19

<210> 5

<211> 10898

<212> DNA  
<213> HUMAN

<220>  
<221> gene  
<222> (1)..(10898)  
<223> full-length MN genomic sequence

<220>  
<221> unsure of base at position 1974  
<222> (1974)  
<223> unsure of base at position 1974, which is in the 5' region flanking the transcription initiation site (3507) as determined by RNase protection assay.

<400> 5  
ggatcctgtt gactcgtgac cttaccccca accctgtgct ctctgaaaca tgagctgtgt 60  
ccactcaggg ttaaatggat taagggcggt gcaagatgtg ctttgtaaa cagatgcttg 120  
aaggcagcat gctcgtaag agtcatcacc aatccctaatt ctcagaataat cagggacaca 180  
aacactgcgg aaggccgcag ggtcctctgc ctaggaaaac cagagacctt tggtcacttg 240  
tttatctgac cttccctcca ctattgtcca tgaccctgcc aaatccccct ctgtgagaaa 300  
cacccaagaa ttatcaataaa aaaaataaaat taaaaaaaaa aataaaaaaa aaaaaaaaaa 360  
aaaaaaaaaa gacttacgaa tagttattga taaatgaata gctattggta aagccaagta 420  
aatgatcata ttcaaaacca gacggccatc atcacagctc aagtctacact gattgtatct 480  
ctttatcatt gtcattctt ggattcacta gattagtcata catcctcaaa attctcccc 540  
aagttctaatt tacgttccaa acattnaggg gttacatgaa gcttgaacct actacctct 600  
ttgctttga gccatgagtt gtaggaatga tgagttaca ctttacatgc tggggattaa 660  
tttaaacttt acctctaagt cagttggta gccttggct tattttgtt gctaattttg 720  
tagttaatgg atgcactgtg aatcttgcta tgatagttt cctccacact ttgccactag 780  
ggtaggttag gtactcagtt ttcagtaatt gcttacctaa gaccctaagc ctttattctc 840  
ttgtactggc ctttatctgt aatatggca tatttaatac aatataattt ttggagttt 900  
tttgggtttt tgggtttttagt acggagtctt gcatctgtca tgcccaggct 960  
ggagtagcag tggtgccatc tcggctact gcaagctcca cttcccgagt tcacgcccatt 1020  
ttcctgcctc agcctcccgaa gtagctggta ctacaggcgc cggccaccat gcccggctaa 1080  
ttttttgtat ttttggtaga gacggggttt caccgttta gccagaatgg tctcgatctc 1140  
ctgacttcgt gatccacccg cctcggcctc ccaaagttct gggattacag gtgtgagcca 1200  
ccgcacctgg ccaattttt gagtcttttta aagtaaaaat atgtcttgcata agctggtaac 1260  
tatggtacat ttccctttttaatgtggta ctgacggta tatagttct tttgagttt 1320  
gcatgcataat gctactttt gcaagtccctt cattacattt ttctctcttc atttgaagag 1380  
catgttatat cttttagctt cacttggctt aaaaggttct ctcattagcc taacacagt 1440  
tcattgtgg taccacttgg atcataagt gaaaaacagt caagaaattt cacagtaata 1500  
cttgggttta agagggatga ttcaaggtaa tctgacacta agaaaactccc ctacctgagg 1560  
tctgagattc ctctgacatt gctgttatata ggctttccct ttgacagcct gtgactgcgg 1620  
actattttc ttaagcaaga tatgctaaag ttttgtgagc cttttccag agagaggtct 1680  
catatctgca tcaagtgaga acatataatg tctgcatgtt tccatatttc aggaatgttt 1740  
gcttgtgttt tatgctttta tatagacagg gaaacttgcctcagtgac ccaaaagagg 1800

tggaaattgt tattggatat catcattggc ccacgcttc tgacccttgg aacaattaag 1860  
ggttcataat ctaaattctg tcagaattgg tacaagaaat agctgttatg tttcttgaca 1920  
ttccacttgg tagaaataa gaatgtgaaa ctcttcagtt ggtgtgtgc cctngtttt 1980  
ttgcaatttc cttcttactg tgtaaaaaaa aagtatgatc ttgctcttag aggtgaggca 2040  
ttcttaatca tcatctttaa agatcaataa tataatcctt tcaaggatta tgtctttatt 2100  
ataataaaaga taatttgtct ttaacagaat caataatata atcccctaaa ggattatatc 2160  
tttgctggc gcagtggctc acacctgtaa tcccagcact ttgggtggcc aaggtggaag 2220  
gatcaaattt gcctacttct atattatctt ctaaagcaga attcatctt cttccctcaa 2280  
tatgtatgata ttgacagggt ttgcctcac tcactagatt gtgagctcct gctcagggca 2340  
ggtagcgtt ttgtttttg ttttgtttt tctttttga gacagggtct tgctctgtca 2400  
cccaggccag agtgcataatgg tacagtctca gctcaactgca gcctcaaccg cctcggctca 2460  
aaccatcatc ccatttcagc ctcctgagta gctgggacta caggcacatg ccattacacc 2520  
tggctaattt ttttgttatt ctatgatgaga cagggtttg ccattgttgc cgggctggc 2580  
tcgaactcct ggactcaagc aatccaccca ctcagcctc cccaaatgag ggaccgtgtc 2640  
ttattcattt ccatgtccct agtccatagc ccagtgcgtt acctatggta gtactaaata 2700  
aatatttgtt gaatgcataa gtaaatagca tttcaggagca caagaacttag attaacaag 2760  
gtggtaaaag gttggagaa aaaaataataa gtttaatttg gcttagatg gagggagagt 2820  
agtaggagac aagatggaaa ggtctcttgg gcaagggtttt gaaggaagtt ggaagtcaga 2880  
agtacacaat gtgcataatcg tggcaggcag tggggagcca atgaaggctt ttgagcagga 2940  
gagtaatgtt ttgaaaaata aatataaggaa aaacctatca gagccccctt gacacatata 3000  
tttgcttttc attcaagctc aagtttgcctt cccacatacc cattacttaa ctcaccctcg 3060  
ggctccccata gcagcctgccc ctacctctt acctgcttcc tgggtggagtc agggatgtat 3120  
acatgagctg cttccctct cagccagagg acatgggggg ccccaagctcc cctgccttcc 3180  
cccttctgtg cctggagctg ggaagcaggc cagggtagc tgaggctggc tggcaagcag 3240  
ctgggtggtg ccagggagag cctgcataatgtt gccaggtgtt gccttgggtt ccaagctagt 3300  
ccatggcccc gataaccttc tgcctgtgca cacacctgccc cctcaactcca ccccccatttc 3360  
agctttggta tgggggagag ggcacaggc cagacaaacc tgtgagactt tggctccatc 3420  
tctgcaaaag ggcgctctgtt gatcagcctt gctccctcc aggcttgctc ctccccacc 3480  
cagctctctgtt ttccaaatgca cgtacagccc gtacacaccg tggctggga caccacacag 3540  
tcagccgcattt ggctccccctg tgccccagcc cctggctccc tctgttgatc ccggccccctg 3600  
ctccaggcctt cactgtgcaat ctgtgtctgtt cactgtctgtt tctgggtgcctt gtccatcccc 3660  
agaggttgcc ccggatgcag gaggattccc cttgggaggagg aggctttctt gggaaagatg 3720  
acccactggg cgaggaggat ctgcccagtg aagaggattt acccagagag gaggatccac 3780  
ccggagagga ggtatctaccc ttggaggaggat atctacctgg agaggaggat ctacctgaag 3840  
ttaagcctaa atcagaagaa gagggtccc tgaagttttaga ggatctaccc actgttgagg 3900  
ctcctggaga tcctcaagaa cccagaata atgcccacag ggacaaagaa ggttaagtgg 3960  
catcaatctc caaatccagg ttccaggagg ttcatgactc ccctccata ccccaagcc 4020  
ggctctgttc actcaggaa ggaggggaga ctgtactccc cacagaagcc cttccagagg 4080  
tcccatacca atatccccat ccccaactctc ggaggttagaa agggacagat gtggagagaa 4140  
aataaaaaagg gtcaaaagg agagaggatgtt gctggatgag atggggagaga agggggaggc 4200  
tggagaagag aaagggatgtt gaactgcaga tgagagaaaa aatgtgcaga cagagaaaa 4260  
aaataggtgg agaaggagag tcagagatgtt tgagggaaaggg agaaaaggaa agcttgggag 4320  
gtgaagtggg taccagagac aagcaagaag agctggatgtt agtcatctca tctttaggct 4380  
caatgaggaa ttgagaccta ggaagaaggg acacagcagg tagagaaacg tggcttctt 4440  
actcccaagc caggaatttg gggaaagggg ttggagacca tacaaggcag agggatgagt 4500  
ggggagaaga aagaaggagaa aaaggaaaga tgggttactc actcattgg gactcaggac 4560

tgaagtgcc actcactttt tttttttttt ttttgagac aaacttcac ttttggcc 4620  
caggctggag tgcaatggcg cgatctcgcc tcactgcaac ctccacccctcc cgggttcaag 4680  
tgattctcct gcctcagcct ctagccaagt agctgcgatt acaggcatgc gccaccacgc 4740  
ccggctaatt tttgtatTTT tagtagagac ggggttgc catgttggc aggctggct 4800  
cgaactcctg atctcaggtg atccaaaccac cctggccctcc caaagtgcg ggattatagg 4860  
cgtgagccac agccctggc ctgaagcagc cactcaCTT tacagaccct aagacaatga 4920  
ttgcaagctg gtaggattgc tgTTTggccc acccagctgc ggtgtttagt ttgggtgcgg 4980  
tctcctgtgc tttgcacctg gccccttaa ggcattgtt acccgtaatg ctccctgtaa 5040  
gcatctgcgt ttgtgacatc gtttggctg ccaggaaggg attggggctc taagctttag 5100  
cggttcatcc ttttcatTT tacagggat gaccagagtc attggcgcta tggaggttag 5160  
acacccaccc gctgcacaga cccaaatctgg gaacccagct ctgtggatct cccctacagc 5220  
cgtccctgaa cactggtccc gggcgtcccc cccgccc accgtcccac cccctcacct 5280  
tttctacccg gttccctaa gttcctgacc taggcgttagt acttcctcac tataactctcc 5340  
caccccgaggc gacccggccct ggccccgggt gtccccagcc tgccggggcc gcttccagtc 5400  
cccggtggat atccggcccc agctcgccgc cttctgcccc gccctgcgccc ccctggaaact 5460  
cctgggcttc cagctcccgc cgctcccaga actgcgcctg cgcaacaatg gccacagtgg 5520  
tgagggggtc tccccggcga gacttggggta tggggggggc cgcaggaaag ggaaccgtcg 5580  
cgcaagtgcct gcccgggggt tgggtggcc ctaccggcg gggccggctc acttgcctct 5640  
ccctacqcaq tgcaactgac cctgcctctt gggcttagaga tggctctggg tcccgggg 5700  
gagtaccggg ctctgcagct gcatctgcac tggggggctg caggtcgcc gggctcgag 5760  
cacactgtgg aaggccaccc tttccctgccc gaggtgagcg cggactggcc gagaaggggc 5820  
aaaggagcgg ggcggacggg ggccagagac gtggccctct cctaccctcg tgccttttc 5880  
agatccacgt ggttccaccc agcaccgcct ttgcacagat tgacgaggcc ttggggcgcc 5940  
egggaggcct ggccgtgtt gcccctttc tggaggtacc agatcctgga cacccctac 6000  
cccccgctt cccatccccat gctccctcccg gactctatcg tggagccaga gacccatcc 6060  
cagcaagctc actcaggccc ctggctgaca aactcattca cgcactgttt gttcatattaa 6120  
caccactgt gaaccaggca ccagccccca acaaggattc tgaagctgta ggtccttgcc 6180  
tctaaggagc ccacagccag tggggggaggc tgacatgaca gacacatagg aaggacatag 6240  
taaagatggg ggtcacagag gaggtgacac ttaaaggcctt cactggtaga aaagaaaaagg 6300  
aggtgttcat tgcaagggaa acagaatgtg caaagactca gaatatggcc tatttaggga 6360  
atggctacat acaccatgtat tagaggaggc ccagtaaagg gaagggatgg tgagatgcct 6420  
gctaggttca ctcactca tttatTTTT tttgacagtc tctctgtcg 6480  
ccaggctgga gtgcagtggt gtgatctgg gtcactgca cttccgcctc cccgggttcaa 6540  
gggattctcc tgccctagct tcctgagtag ctggggttac aggtgtgtgc caccatgccc 6600  
agctaatttt ttttggatt ttttagtagac agggttcac catgttggc aggctggct 6660  
caaactcctg gcctcaagtg atccgcctga ctcagcctac caaagtgcgt attacaagt 6720  
tgagccaccg tgcccagcca cactcaCTG ttcttaatg ccagccacac agcacaaaagt 6780  
tcagagaaat gcctccatca tagcatgtca atatgttcat actcttaggt tcatgatgtt 6840  
cttaacatta ggttcataag caaaataaga aaaaagaata ataaataaaa gaagtggcat 6900  
gtcaggaccc cacctgaaaaa gccaaacaca gaatcatgaa ggtgaatgca gaggtgacac 6960  
caacacaaaag gtgttatata ggtttcctgt ggggagttatg tacggaggca gcagtggatg 7020  
agactgcaaa cgtcagaagg gcacgggtca ctgagagccct agtacccatg taaagtgggc 7080  
tctctccctc tctctccagc ttgtcattga aaaccagtc accaagctt gttgggtcgca 7140  
cagcaagagt acatagagtt tgaaataata cataggattt taagagggag acactgtctc 7200  
taaaaaaaaaa aacaacagca acaacaaaaa gcaacaacca ttacaatttt atgttccctc 7260  
agcattctca gagctgagga atgggagagg actatggaa ccccttcat gttccggcct 7320

tcagccatgg ccctggatac atgcactcat ctgtcttaca atgtcattcc cccaggaggg 7380  
cccggaaagaa aacagtgcct atgagcagtt gctgtctcg tcgctgagga 7440  
agtcagtt gtttgtctgg ccactaatct ctgtggccta gttcataaag aatcacccctt 7500  
tggagcttca ggtctgaggg tggagatggg ctccctccag tgccaggaggg attgaagcat 7560  
gagccagcgc tcacatgt aataaccatg aagctgacag acacagttac ccgcaaacgg 7620  
ctgcctacag attaaaaacc aagcaaaaac cgccgggcac ggtggctcac gcctgttaatc 7680  
ccagcactt gggaggccaa ggcaggtgga tcacgaggc aagagatcaa gaccatcctg 7740  
gccaacatgg tgaaacccca tctctactaa aaatacgaaa aaatagccag gcgtggtggc 7800  
gggtgcctgt aatcccagct actcgggagg ctgaggcagg agaatggcat gaacccggga 7860  
ggcagaagtt gcagtgagcc gagatcgtgc cactgcactc cagcctggc aacagagcga 7920  
gactcttgc tcaaaaaaaaaaaaaaaa gaaaaccaag caaaaaaccaa aatgagacaa 7980  
aaaaaaacaag accaaaaaaat ggtgtttgga aattgtcaag gtcaagtctg gagagctaaa 8040  
cttttctga gaactgttta tcttaataa gcatcaaata tttaacttt gtaaataactt 8100  
ttgttgaaaa tcgttctctt cttagtcact cttgggtcat tttaatctc acttactcta 8160  
ctagacctt taggttctg cttagactagg tagaactctg ctttgcatt tcttgtgtct 8220  
gttttgtata gttatcaata ttcatattta ttacaagtt attcagatca tttttcttt 8280  
tctttttttt tttttttttt ttttttacat ctttagtaga gacagggttt caccatattg 8340  
gccaggctgc tctcaaactc ctgaccttgc gatccaccag cctcggcctc ccaaagtgc 8400  
gggattcatt ttttctttt aatttgcgtt gggcttaaac ttgtggccca gcactttatg 8460  
atggtacaca gagttaaagag tgttagactca gacggcttt cttcttcct tctcttcctt 8520  
cctcccttcc ctcccacctt cccttctctc cttcccttct ttcttcctct cttgcttcct 8580  
caggcctttt ccagttgc tcaagccctg tactttttt tgagttaaacg tcttatggga 8640  
agggcctgca ctttagtgaag aagtggctc agagttgagt taccttggct tctgggaggt 8700  
gaaactgtat ccctataaccc tgaagcttta agggggtgca atgttagatga gaccccaaca 8760  
tagatcctt tcacaggctc agagactcag gtcccaggac tggacatatc tgcactcctg 8820  
ccctctgact tcagccgcta cttccaatat gaggggtctc tgactacacc gccctgtgcc 8880  
cagggtgtca tctggactgt tttaaccag acagtgtatgc tgagtgtctaa gcaggtggc 8940  
etgggggtgtg tggacaca gtgggtgcgg gggaaagagg atgtaagatg agatgagaaaa 9000  
caggagaaga aagaaatcaa ggctgggctc tgtggcttac gcctataatc ccaccacgtt 9060  
gggaggctga ggtgggagaa tggttttagc ccaggagttc aagacaaggc gggcaacat 9120  
agtgtgaccc catctctacc aaaaaaaaaacc caacaaaacc aaaaatagcc gggcatggtg 9180  
gtatgcggcc tagtcccagc tactcaagga ggctgagggtg ggaagatcgc ttgattccag 9240  
gagtttgaga ctgcagttag ctatgatccc accactgcct accatctta ggatacattt 9300  
atttattttt aaaaagaaatc aagaggctgg atgggaata caggagctgg agggtggagc 9360  
cctgaggtgc tgggtgttag ctggcctggg acccttgcct cctgtcatgc catgaaccca 9420  
cccacactgt ccactgaccc cttagctcc acaccctctc tgacaccctg tggggacctg 9480  
gtgactctg gctacagctg aacctccgag cgacgcagcc tttgaatggg cgagtgattg 9540  
aggcctcctt ccctgctgga gtggacagca gtcctcgcc tgctgagcca ggtacagctt 9600  
tgtctggttt cccccccagcc agtagtccct tatcctccca tgtgtgtgcc agtgtctgtc 9660  
attgggtgtc acagccccc tctcacatct ctttttctc tccagtcagcc ctgaattcct 9720  
gcctggctgc tggtagtct gcccctcctc ttggccttgc tgccaggaga ctccctcagca 9780  
ccattcagcc ccagggtgtc tcaggaccgc ctctgctccc tctcccttcc tgcagaacag 9840  
accccaaccc caatattaga gaggcagatc atgggggaa ttccccatt gtccccagag 9900  
gctaattgat tagaatgaag cttgagaaat ctcccagcat ccctctcgca aaagaatccc 9960  
ccccccctttt tttaaagata gggctctact ctgtttgccc caggctgggg tgggtgtggca 10020  
cgatcatagc tcactgcagc ctgcactcc taggctcagg caatccttc accttagctt 10080

ctcaaagcac tgggactgta ggcattgagcc actgtgcctg gccccaaacg gccctttac 10140  
ttggcttta ggaagcaaaa acggtgctta tcttacccct tctcgtgtat ccaccctcat 10200  
ccctggctg gcctcttctg gagactgagg cactatgggg ctgcctgaga actcggggca 10260  
gggtggtgg agtgcactga ggcaggtgtt gaggaactct gcagacccct cttccttccc 10320  
aaagcagccc tctctgctct ccatcgagg tgacatccta gccctgggtt ttggcctcct 10380  
tttgctgtc accagcgtcg cgttccttgt gcagatgaga aggccagcaca ggtattacac 10440  
tgacccttc ttcaaggcaca agcttcccc acccttgtgg agtcacttca tgcaaagcgc 10500  
atgcaaataa gctgctcctg ggccagttt ctgattagcc tttcctgttg tgtacacaca 10560  
gaaggggaaac caaagggggt gtgagctacc gcccagcaga gtagccgag actggagcct 10620  
agaggctgga tcttggagaa tgtgagaagc cagccagagg catctgaggg ggagccggta 10680  
actgtcctgt cctgctcatt atgccacttc ctttaactg ccaagaaatt tttaaaata 10740  
aatatttata ataaaatatg tgtagtcac ctttggccc caaatcagaa ggaggtattt 10800  
gaatttccta ttactgttat tagcaccaat tttagtggtaa tgcatttatt ctattacagt 10860  
tcggcctcct tccacacatc actccaatgt gttgctcc 10898

<210> 6  
<211> 37  
<212> PRT  
<213> HUMAN

<400> 6  
Met Ala Pro Leu Cys Pro Ser Pro Trp Leu Pro Leu Leu Ile Pro Ala  
1 5 10 15  
Pro Ala Pro Gly Leu Thr Val Gln Leu Leu Leu Ser Leu Leu Leu  
20 25 30  
Met Pro Val His Pro  
35

<210> 7  
<211> 25  
<212> DNA  
<213> HUMAN

<400> 7  
tggggttctt gaggatctcc aggag 25

<210> 8  
<211> 26  
<212> DNA  
<213> HUMAN

<400> 8  
ctctaacttc agggagccct cttctt

26

<210> 9  
<211> 48  
<212> DNA  
<213> HUMAN

<220>  
<221> primer\_bind  
<222> (1)..(48)  
<223> anchor primer that anneals to the homopolymeric tail.

<220>  
<221> inosine  
<222> (36)..(37) (41)..(42) (46)..(47)  
<223> each of the modified\_bases at positions (36), (37), (41), (42), (46) and (47) are inosine

<400> 9  
cuacuacuac uaggccacgc gtcgactagt acgggaaggg aaggaaag 48

<210> 10  
<211> 6  
<212> PRT  
<213> HUMAN

<400> 10  
Glu Glu Asp Leu Pro Ser  
1 5

<210> 11  
<211> 6  
<212> PRT  
<213> HUMAN

<400> 11  
Gly Glu Asp Asp Pro Leu  
1 5

<210> 12  
<211> 21

<212> PRT  
<213> HUMAN

<400> 12  
Asn Asn Ala His Arg Asp Lys Glu Gly Asp Asp Gln Ser His Trp Arg  
1 5 10 15  
Tyr Gly Gly Asp Pro  
20

<210> 13  
<211> 16  
<212> PRT  
<213> HUMAN

<400> 13  
His Pro Gln Arg Leu Pro Arg Met Gln Glu Asp Ser Pro Leu Gly Gly  
1 5 10 15

<210> 14  
<211> 24  
<212> PRT  
<213> HUMAN

<400> 14  
Glu Glu Asp Ser Pro Arg Glu Glu Asp Pro Pro Gly Glu Glu Asp Leu  
1 5 10 15

Pro Gly Glu Glu Asp Leu Pro Gly  
20

<210> 15  
<211> 13  
<212> PRT  
<213> HUMAN

<400> 15  
Leu Glu Glu Gly Pro Glu Glu Asn Ser Ala Tyr Glu Gln  
1 5 10

<210> 16  
<211> 16

<212> PRT  
<213> HUMAN

<400> 16  
Met Arg Arg Gln His Arg Arg Gly Thr Lys Gly Gly Val Ser Tyr Arg  
1 5 10 15

<210> 17  
<211> 45  
<212> DNA  
<213> HUMAN

<400> 17  
gtcgcttagct ccatgggtca tatgcagagg ttgccccgga tgcag 45

<210> 18  
<211> 43  
<212> DNA  
<213> HUMAN

<400> 18  
gaagatctct tactcgagca ttctccaaga tccagcctct agg 43

<210> 19  
<211> 10  
<212> DNA  
<213> HUMAN

<400> 19  
ctccatctct 10

<210> 20  
<211> 10  
<212> DNA  
<213> HUMAN

<400> 20  
ccaccccccatt 10

<210> 21

<400> 21

000

<210> 22

<211> 8

<212> PRT

<213> HUMAN

<400> 22

Leu Glu His His His His His His

1 5

<210> 23

<211> 10

<212> DNA

<213> HUMAN

<220>

<221> misc\_feature

<222> (1)..(10)

<400> 23

Yyyccayyyyy

10

<210> 24

<211> 10

<212> DNA

<213> HUMAN

<300>

<301> Locker and Buzard,

<303> DNA Sequencing and Mapping

<304> 1

<306> 3-11

<307> 1990

<400> 24

tgtgagactt

10

<210> 25

<211> 4

<212> PRT

<213> HUMAN

<220>  
<221> SITE  
<222> (1)..(4)  
<223> sequence element defined by Suzuki, J. Mol. Biol., 207: 61-84 (1989)  
as motif frequently found in gene regulatory proteins.

<220>  
<221> VARIANTS  
<222> (3)..(4)  
<223> variants in sequence element defined by Suzuki, J. Mol. Biol., 207:  
61-84 (1989) as motif frequently found in gene regulatory proteins.

<400> 25  
Ser Pro Xaa Xaa

1

<210> 26  
<211> 4  
<212> PRT  
<213> HUMAN

<220>  
<221> SITE  
<222> (1)..(4)  
<223> sequence element defined by Suzuki, J. Mol. Biol., 207: 61-84 (1989)  
as a motif frequently found in gene regulatory proteins.

<220>  
<221> VARIANTS  
<222> (3)..(4)  
<223> variants in sequence element defined by Suzuki, J. Mol. Biol., 207:  
61-84 (1989) as a motif frequently found in gene regulatory proteins.

<400> 26  
Thr Pro Xaa Xaa  
1

<210> 27  
<211> 540  
<212> DNA  
<213> HUMAN

<220>

<221> promoter  
<222> (1) .. (540)

<400> 27  
cttgctttc attcaagctc aagttgtct cccacatacc cattactaa ctcaccctcg 60  
ggctccccta gcagcctgcc ctacctctt acctgcttcc tggggagtc agggatgtat 120  
acatgagctg cttccctct cagccagagg acatgggggg ccccagctcc cctgcctttc 180  
cccttctgtg cctggagctg ggaagcagggc cagggtagc tgaggctggc tggcaagcag 240  
ctgggtggtg ccagggagag cctgcatagt gccaggttgt gccttgggtt ccaagctagt 300  
ccatggccccc gataaaccttc tgcctgtgca cacacctgcc cctcaactcca cccccatcct 360  
agctttggta tgggggagag ggcacagggc cagacaaacc tgtgagactt tggctccatc 420  
tctgcaaaag ggcgctctgt gagtcagcct gctccctcc aggcttgctc ctccccacc 480  
cagctctcgt ttccaatgca cgtacagccc gtacacacccg tgtgctggta caccacacag 540

<210> 28  
<211> 445  
<212> DNA  
<213> HUMAN

<220>  
<221> exon  
<222> (1)  
<223> 1st MN exon

<400> 28  
ccccgtacac accgtgtgct gggacacccc acagtcagcc gcatggctcc cctgtgcccc 60  
agccccctggc tcctctgtt gatcccgcc cctgctccag gcctcaactgt gcaactgctg 120  
ctgtcaactgc tgcttctggc gcctgtccat ccccaagaggt tgcccccggat gcaggaggat 180  
tcccccttgg gaggaggctc ttctgggaa gatgacccac tggcgagga ggatctgccc 240  
agtgaagagg attcacccag agaggaggat ccacccggag aggaggatct acctggagag 300  
gaggatctac ctggagagga ggatctacct gaagttaaagc ctaaatcaga agaagagggc 360  
tccctgaagt tagaggatct acctactgtt gaggctcctg gagatcctca agaacccag 420  
aataatgccc acagggacaa agaag 445

<210> 29  
<211> 30  
<212> DNA  
<213> HUMAN

<220>  
<221> exon  
<222> (1)  
<223> 2nd MN exon

<400> 29  
gggatgacca gagtcattgg cgctatggag 30

<210> 30  
<211> 171  
<212> DNA  
<213> HUMAN

<220>  
<221> exon  
<222> (1)  
<223> 3rd MN exon

<400> 30  
gcgcacccgcc ctggccccgg gtgtccccag cctgcgcggg ccgccttccag tccccgggtgg 60  
atatccgccc ccagctcgcc gccttctgcc cggccctgcg ccccccggaa ctccctgggct 120  
tccagctccc gccgctccca gaactgcgc g tgcgcaacaa tggccacagt g 171

<210> 31  
<211> 143  
<212> DNA  
<213> HUMAN

<220>  
<221> exon  
<222> (1)  
<223> 4th MN exon

<400> 31  
tgcaactgac cctgcctcct gggcttagaga tggctctggg tccccggcgg gagtaccggg 60  
ctctgcagct gcatctgcac tggggggctg caggtagtcc gggctcggag cacactgtgg 120  
aaggccaccc ttccctgcc gag 143

<210> 32  
<211> 93  
<212> DNA  
<213> HUMAN

<220>  
<221> exon  
<222> (1)  
<223> 5th MN exon

<400> 32  
atccacgtgg ttcacacctag caccgcctt gccagagttg acgaggcctt ggggcgcaccg 60  
ggaggcctgg ccgtgttggc cgcccttctg gag 93

<210> 33  
<211> 67  
<212> DNA  
<213> HUMAN

<220>  
<221> exon  
<222> (1)  
<223> 6th MN exon

<400> 33  
gagggcccg aagaaaacag tgcctatgag cagttgctgt ctcgcttgga agaaatcgct 60  
gaggaag 67

<210> 34  
<211> 158  
<212> DNA  
<213> HUMAN

<220>  
<221> exon  
<222> (1)  
<223> 7th MN exon

<400> 34  
gctcagagac tcaggtccca ggactggaca tatctgcact cctgcctct gacttcagcc 60  
gctacttcca atatgagggg tctctgacta caccgcctg tgcccagggt gtcatctgga 120  
ctgtgttaa ccagacagtg atgctgagtg ctaagcag 158

<210> 35  
<211> 145  
<212> DNA  
<213> HUMAN

<220>  
<221> exon  
<222> (1)  
<223> 8th MN exon

<400> 35  
ctccacaccc tctctgacac cctgtggga cctggtgact ctcggctaca gctgaacttc 60  
cgagcgacgc agccttgaa tggcgagtg attgaggcct cttccctgc tggagtggac 120  
agcagtcctc gggctgctga gccag 145

<210> 36  
<211> 27  
<212> DNA  
<213> HUMAN

<220>  
<221> exon  
<222> (1)  
<223> 9th MN exon

<400> 36  
tccagctgaa ttccctgcctg gctgctg 27

<210> 37  
<211> 82  
<212> DNA  
<213> HUMAN

<220>  
<221> exon  
<222> (1)  
<223> 10th MN exon

<400> 37  
gtgacatcct agccctgggtt tttggcctcc ttttgctgt caccagcgtc gcgttccttg 60  
tgcagatgag aaggcagcac ag 82

<210> 38  
<211> 191  
<212> DNA  
<213> HUMAN

<220>  
<221> exon  
<222> (1)  
<223> 11th MN exon

<400> 38

aaggggaaacc aaaggggggtg tgagctaccg cccagcagag gtagccgaga ctggagccta 60  
gaggctggat cttggagaat gtgagaagcc agccagaggc atctgagggg gagccggtaa 120  
ctgtcctgtc ctgctcatta tgccacttcc ttttaactgc caagaaaattt tttaaaataa 180  
atatttataa t 191

<210> 39

<211> 1174

<212> DNA

<213> HUMAN

<220>

<221> intron

<222> (1)..(1174)

<223> 1st MN intron

<400> 39

gtaagtggtc atcaatctcc aaatccaggt tccaggaggt tcatacgactcc cctcccatac 60  
cccagcctag gctctgttca ctcaggaaag gaggggagac tgtactcccc acagaagccc 120  
ttccagaggt cccataccaa tatccccatc cccactctcg gaggtagaaa gggacagatg 180  
tggagagaaa ataaaaaggg tgcaaaagga gagaggtgag ctggatgaga tggagagaaa 240  
gggggaggct ggagaagaga aaggatgag aactgcagat gagaaaaaaa atgtgcagac 300  
agagaaaaaa aataggtgga gaaggagagt cagagagttt gaggggaaga gaaaaggaaa 360  
gcttgggagg tgaagtgggt accagagaca agcaagaaga gctggtagaa gtcatctcat 420  
cttaggctac aatgaggaat tgagacctag gaagaaggga cacagcaggt agagaaacgt 480  
ggcttcttga ctcccaagcc aggaatttgg ggaaagggtt tggagaccat acaaggcaga 540  
ggatgagtg gggagaagaa agaaggaga aaggaaagat ggtgtactca ctcatttggg 600  
actcaggact gaagtgccca ctcactttt tttttttt ttttgagaca aactttcact 660  
tttgttgcgg aggctggagt gcaatggcgc gatctcggt cactgcaacc tccacccccc 720  
gggttcaagt gattctcctg cctcagcctc tagccaagta gctgcgatta caggcatgcg 780  
ccaccacgcc cgcttaattt ttgtatTTT agtagagacg gggttgcggcc atgttggtca 840  
ggctggtctc gaactcctga tctcagggtga tccaaaccacc ctggcctccc aaagtgcgg 900  
gattataggc gtgagccaca gcgcctggcc tgaagcagcc actcactttt acagacccta 960  
agacaatgtat tgcaagctgg taggattgct gtttggccca cccagctgcg gtgttgagtt 1020  
tgggtgcggc ctccctgtgct ttgcacctgg cccgcttaag gcatttgcgc cccgtaatgc 1080  
tcctgttaagg catctgcgtt tgtgacatcg ttttggtcgc caggaaggaa ttggggctct 1140  
aagcttgagc ggttcatcct tttcatttat acag 1174

<210> 40

<211> 193

<212> DNA

<213> HUMAN

<220>

<221> intron

<222> (1)..(193)  
<223> 2nd MN intron

<400> 40  
gtgagacacc caccgctgc acagacccaa tctggaaacc cagctctgtg gatctccct 60  
acagccgtcc ctgaacactg gtcccgccg tcccacccgc cgcccaccgt cccacccct 120  
cacctttct acccggttc cctaagttcc tgacctaggc gtcagacttc ctcactatac 180  
tctcccaccc cag 193

<210> 41  
<211> 131  
<212> DNA  
<213> HUMAN

<220>  
<221> intron  
<222> (1)..(131)  
<223> 3rd MN intron

<400> 41  
gtgagggggt ctccccgccc agacttgggg atggggcgaa ggcgcaggaa ggaaaccgtc 60  
gcccgggtt tgcccgaaaa ttgggctggc cttaccgggc gggccggct cacttgcc 120  
ccctacgca g 131

<210> 42  
<211> 89  
<212> DNA  
<213> HUMAN

<220>  
<221> intron  
<222> (1)..(89)  
<223> 4th MN intron

<400> 42  
gtgagcgccg actggccgag aaggggcaaa ggagcgccgc ggacggggc cagagacgtg 60  
gccctctcct accctcggt cctttcag 89

<210> 43  
<211> 1400  
<212> DNA  
<213> HUMAN

<220>  
<221> intron  
<222> (1)..(1400)  
<223> 5th MN intron

<400> 43  
gtaccagatc ctggacaccc cctactcccc gctttccat cccatgctcc tcccggactc 60  
tatcgtaggag ccagagaccc catcccagca agctcaactca ggcccctggc tgacaaaactc 120  
attcacgcac tggttgttca tttaacaccc actgtgaacc aggccaccagc ccccaacaag 180  
gattctgaag ctgttaggtcc ttgcctctaa ggagcccaca gccagtgggg gaggctgaca 240  
tgacagacac ataggaagga catagtaaag atgggttgtca cagaggaggt gacacttaaa 300  
gccttcactg gttagaaaaga aaaggaggtg ttcattgcag aggaaacaga atgtgcaaag 360  
actcagaata tggcctattt agggaatggc tacatacacc atgatttagag gaggcccagt 420  
aaagggaaagg gatggtgaga tgcctgctag gttcaactcac tcacttttat ttattttattt 480  
attttttga cagtcctctt gtcgcccagg ctggagtgca gtgggtgtat cttgggtcac 540  
tgcaacttcc gcctcccggg ttcaagggat ttcctgcct cagttcctg agtagctggg 600  
gttacaggtg tggccacca tgcccagcta atttttttt gtatttttag tagacagggt 660  
ttcaccatgt tggtcaggct ggtctaaac tcctggcctc aagtgtatccg cctgactcag 720  
cctaccaaag tgctgattac aagtgtgagc caccgtgccc agccacactc actgattctt 780  
taatgccagc cacacagcac aaagttcaga gaaatgcctc catcatagca tgtcaatatg 840  
ttcatactct taggttcatg atgttcttaa cattaggttc ataagaaaaa taagaaaaaaa 900  
gaataataaa taaaagaagt ggcatgtcag gacctcacct gaaaagccaa acacagaatc 960  
atgaaggta atgcagaggt gacaccaaca caaagggtta tatatggttt cctgtgggga 1020  
gtatgtacgg aggcagcagt gagtgagact gcaaacgtca gaagggcacg ggtcaactgag 1080  
agcctagtat cctagtaaaag tggctctct ccctctctt ccagcttgc attgaaaaacc 1140  
agtccaccaa gctgttggt tcgcacagca agagtacata gagttgaaa taatacatag 1200  
gatttaaga gggagacact gtctctaaaa aaaaaaacaa cagcaacaac aaaaagcaac 1260  
aaccattaca attttatgtt ccctcagcat tctcagagct gaggaatggg agaggactat 1320  
ggaaacccccc ttcatgttcc ggccttcagc catggccctg gatacatgca ctcatctgtc 1380  
ttacaatgtc attccccccag 1400

<210> 44  
<211> 1334  
<212> DNA  
<213> HUMAN

<220>  
<221> intron  
<222> (1)..(1334)  
<223> 6th MN intron

<400> 44  
gtcagttgt tggctggcc actaatctct gtggcctagt tcataaagaa tcacccttg 60  
gagcttcagg tctgaggctg gagatggct ccctccagtg caggaggat tgaagcatga 120  
gccagcgctc atcttgataa taaccatgaa gctgacagac acagttaccc gcaaacggct 180

gcctacagat tgaaaaccaa gcaaaaaccg ccgggcacgg tggctcacgc ctgtaatccc 240  
agactttgg gaggccaagg cagggtgatc acgaggtaa gagatcaaga ccatcctggc 300  
caacatggtg aaaccccatc tctactaaaa atacaaaaaa atagccaggc gtggtggcgg 360  
gtgcctgtaa tcccagctac tcgggaggct gaggcaggag aatggcatga acccgggagg 420  
cagaagttgc agtgagccga gatcggtcca ctgcactcca gcctggcaa cagagcgaga 480  
ctcttgtctc aaaaaaaaaa aaaaaaaaaa aaaccaagca aaaaccaaaa tgagacaaaa 540  
aaaacaagac caaaaaatgg tgtttgaaa ttgtcaaggt caagtctgga gagctaaact 600  
tttctgaga actgtttatc ttataataagc atcaaataatt ttaacttgt aaatacttt 660  
gttggaaatc gttctcttct tagtcactt tgggtcattt taaatctcac ttactctact 720  
agaccttttta ggtttctgct agacttagta gaactctgcc tttgcatttc ttgtgtctgt 780  
tttgtatagt tatcaatatt catattatt tacaagttt tcagatcatt ttttctttt 840  
ttttttttt tttttttttt ttttacatct ttagtagaga cagggttca ccatattggc 900  
caggctgctc tcaaactcct gaccttgtga tccaccagcc tcggcctccc aaagtgctgg 960  
gattcatttt ttcttttaa tttgtctgg gcttaaactt gtggcccaagc actttatgtat 1020  
ggtacacaga gttaagagtg tagactcaga cggtctttct tctttccttc tcttccttcc 1080  
tcccttcctt cccaccttcc cttctctcct tcccttctt cttctctct tgcttcctca 1140  
ggcctttcc agttgctcca aagccctgta ctttttttg agttaacgtc ttatggaaag 1200  
ggcctgcact tagtgaagaa gtggtctcag agttgagttt ccttggcttc tgggaggtga 1260  
aactgtatcc ctataccctg aagcttaag ggggtgcaat gtagatgaga ccccaacata 1320  
gatcctcttc acag 1334

<210> 45  
<211> 512  
<212> DNA  
<213> HUMAN

<220>  
<221> intron  
<222> (1)..(512)  
<223> 7th MN intron

<400> 45  
gtgggcctgg ggtgtgtgtg gacacagtgg gtgcggggqa aagaggatgt aagatgagat 60  
gagaaacagg agaagaaaga aatcaaggct gggctctgtg gcttacgcct ataatcccac 120  
cacgttggga ggctgagggtg ggagaatggt ttgagccccag gagttcaaga caaggcgggg 180  
caacatagtg tgacccatc tctacaaaaa aaaccccaac aaaaccaaaa atagccgggc 240  
atggtggtat gccccctagt cccagctact caaggaggct gaggtggaa gatcgcttga 300  
ttccaggagt ttgagactgc agttagctat gatcccacca ctgcctacca tctttaggat 360  
acatttattt attataaaaa gaaatcaaga ggctggatgg ggaatacagg agctggaggg 420  
tggagccctg aggtgctgg tggtagctgg cctgggaccc ttgtttcctg tcatgccatg 480  
aacccaccca cactgtccac tgacccctt ag 512

<210> 46  
<211> 114

<212> DNA  
<213> HUMAN

<220>  
<221> intron  
<222> (1)..(114)  
<223> 8th MN intron

<400> 46  
gtacagctt gtctggtttc cccccagcca gtagtccctt atcctccat gtgtgtgcca 60  
gtgtctgtca ttggtggtca cagcccgct ctcacatctc cttttctct ccag 114

<210> 47  
<211> 617  
<212> DNA  
<213> HUMAN

<220>  
<221> intron  
<222> (1)..(617)  
<223> 9th MN intron

<400> 47  
gtgagtctgc ccctcctctt ggtcctgatg ccaggagact cctcagcacc attcagcccc 60  
agggctgctc aggaccgcct ctgctccctc tcctttctg cagaacagac cccaacccca 120  
atattagaga ggcagatcat ggtggggatt cccccattgt ccccagaggc taattgatta 180  
gaatgaagct tgagaaatct cccagcatcc ctctcgcaaa agaatcccccc ccccttttt 240  
taaagatagg gtctcactct gtttgcucca ggctgggtg ttgtggcacg atcatagctc 300  
actgcagcct cgaactccta ggctcaggca atcctttcac cttagcttct caaagcactg 360  
ggactgttagg catgagccac tgtgcctggc cccaaacggc cctttactt ggcttttagg 420  
aagcaaaaac ggtgcttatac ttacccttc tcgtgtatcc accctcatcc cttggctggc 480  
ctcttctgga gactgaggca ctatggggct gcctgagaac tcggggcagg ggtgggtggag 540  
tgcactgagg caggtgttga ggaactctgc agaccctct tccttccaa agcagccctc 600  
tctgctctcc atcgcag 617

<210> 48  
<211> 130  
<212> DNA  
<213> HUMAN

<220>  
<221> intron  
<222> (1)..(130)  
<223> 10th MN intron

<400> 48  
gtattacact gaccctttct tcaggcacaa gcttccccca cccttgga gtcacttcat 60  
gcaaaggcga tgcaaatgag ctgctcctgg gccagtttc tgattagcct ttcctgttgt 120  
gtacacacag 130

<210> 49  
<211> 1401  
<212> DNA  
<213> HUMAN

<400> 49  
caaacttca ctttgttgc ccaggctgga gtgcaatggc gcgatctcg 60  
cctccaccc cggggttcaa gtgattctcc tgcctcagcc tctagccaag tagctgcgat 120  
tacaggcatg cgccaccacg cccggtaat tttgtattt ttagtagaga cggggtttcg 180  
ccatgttggt caggctggtc tcgaactcct gatctcaggt gatccaacca ccctggcctc 240  
ccaaagtgt gggattata ggcgtgagcca cagcgcctgg cctgaagcag ccactcactt 300  
ttacagaccc taagacaatg attgcaagct ggtaggattt ctgtttggcc cacccagctg 360  
gggtgtttagt tttgggtgcg gtctcctgtg ctttgcacct ggcccgtta aggcatttgt 420  
tacccgtaat gctcctgtaa ggcacatcg 60  
gattggggct ctaagcttga gcgggttcatc ctttcattt atacagggga tgaccagagt 540  
cattggcgct atggaggtga gacacccacc cgctgcacag acccaatctg ggaacccagc 600  
tctgtggatc tcccctacag ccgtccctga acactggtcc cggcggtccc acccgccgccc 660  
caccgtccca ccccttcacc ttttctaccc gggttcccta agttcctgac ctaggcgtca 720  
gacttcctca ctatactctc ccaccccagg cgacccgccc tggcccccggg tgtccccagc 780  
ctgcgcgggc cgcttccagt ccccggtgga tatccggccc cagctcgccg ccttctgccc 840  
ggccctgcgc cccctggAAC tcctgggctt ccagctcccg ccgctcccag aactgcgcct 900  
gcgcacaat ggccacagtg gtgagggggt ctcccccgg agacttgggg atggggcg 960  
gcgcaggaa gggAACCGTC gcgcagtggc tgcccggggg ttgggctggc cctaccggc 1020  
ggggccggct cacttgcctc tccctacgca gtgcaactga ccctgcctcc tggctagag 1080  
atggctctgg gtcccgggcg ggagtaccgg gctctgcagc tgcatctgca ctgggggct 1140  
gcaggtcgTC cgggctcgGA gcacactgtg gaaggccacc gttccctgc cgaggtgagc 1200  
gcggactggc cgagaagggg caaaggagcg gggcggacgg gggccagaga cgtggccctc 1260  
tcctaccctc gtgtccttt cagatccacg tggttcacct cagcaccgccc tttgccagag 1320  
ttgacgaggc cttggggcgc ccgggaggcc tggccgtgtt ggccgcctt ctggaggtaC 1380  
cagatcctgg acacccctca c 1401

<210> 50  
<211> 59  
<212> PRT  
<213> HUMAN

<400> 50  
Ser Ser Gly Glu Asp Asp Pro Leu Gly Glu Glu Asp Leu Pro Ser Glu  
1 5 10 15

Glu Asp Ser Pro Arg Glu Glu Asp Pro Pro Gly Glu Glu Asp Leu Pro  
20 25 30

Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Glu Val Lys Pro  
35 40 45

Lys Ser Glu Glu Glu Gly Ser Leu Lys Leu Glu  
50 55

<210> 51

<211> 257

<212> PRT

<213> HUMAN

<400> 51

Gly Asp Asp Gln Ser His Trp Arg Tyr Gly Gly Asp Pro Pro Trp Pro  
1 5 10 15

Arg Val Ser Pro Ala Cys Ala Gly Arg Phe Gln Ser Pro Val Asp Ile  
20 25 30

Arg Pro Gln Leu Ala Ala Phe Cys Pro Ala Leu Arg Pro Leu Glu Leu  
35 40 45

Leu Gly Phe Gln Leu Pro Pro Leu Pro Glu Leu Arg Leu Arg Asn Asn  
50 55 60

Gly His Ser Val Gln Leu Thr Leu Pro Pro Gly Leu Glu Met Ala Leu  
65 70 75 80

Gly Pro Gly Arg Glu Tyr Arg Ala Leu Gln Leu His Leu His Trp Gly  
85 90 95

Ala Ala Gly Arg Pro Gly Ser Glu His Thr Val Glu Gly His Arg Phe  
100 105 110

Pro Ala Glu Ile His Val Val His Leu Ser Thr Ala Phe Ala Arg Val  
115 120 125

Asp Glu Ala Leu Gly Arg Pro Gly Gly Leu Ala Val Leu Ala Ala Phe  
130 135 140

Leu Glu Glu Gly Pro Glu Glu Asn Ser Ala Tyr Glu Gln Leu Leu Ser  
145 150 155 160

Arg Leu Glu Glu Ile Ala Glu Glu Gly Ser Glu Thr Gln Val Pro Gly  
165 170 175

Leu Asp Ile Ser Ala Leu Leu Pro Ser Asp Phe Ser Arg Tyr Phe Gln  
180 185 190

Tyr Glu Gly Ser Leu Thr Thr Pro Pro Cys Ala Gln Gly Val Ile Trp  
195 200 205

Thr Val Phe Asn Gln Thr Val Met Leu Ser Ala Lys Gln Leu His Thr  
210 215 220

Leu Ser Asp Thr Leu Trp Gly Pro Gly Asp Ser Arg Leu Gln Leu Asn  
225 230 235 240

Phe Arg Ala Thr Gln Pro Leu Asn Gly Arg Val Ile Glu Ala Ser Phe  
245 250 255

Pro

<210> 52

<211> 20

<212> PRT

<213> HUMAN

<400> 52

Ile Leu Ala Leu Val Phe Gly Leu Leu Phe Ala Val Thr Ser Val Ala  
1 5 10 15

Phe Leu Val Gln  
20

<210> 53

<211> 25

<212> PRT

<213> HUMAN

<400> 53

Met Arg Arg Gln His Arg Arg Gly Thr Lys Gly Gly Val Ser Tyr Arg  
1 5 10 15

Pro Ala Glu Val Ala Glu Thr Gly Ala  
20 25

<210> 54  
<211> 59  
<212> PRT  
<213> HUMAN

<400> 54  
Ser Ala Ser Glu Glu Pro Ser Pro Ser Glu Val Pro Phe Pro Ser Glu  
1 5 10 15

Glu Pro Ser Pro Ser Glu Glu Pro Phe Pro Ser Val Arg Pro Phe Pro  
20 25 30

Ser Val Val Leu Phe Pro Ser Glu Glu Pro Phe Pro Ser Lys Glu Pro  
35 40 45

Ser Pro Ser Glu Glu Pro Ser Ala Ser Glu Glu  
50 55

<210> 55  
<211> 470  
<212> RNA  
<213> HUMAN

<400> 55  
cauggcccg auaacuuucu gccugugcac acaccugccc cucacuccac ccccauccua 60  
gcuuugguai gggggagagg gcacagggcc agacaaaccu gugagacuuu ggcuccaucu 120  
cugcaaaagg gcgcucugug agucagccug cucccccucca ggcuugcucc ucccccaccc 180  
agcucucguu uccaaugcac guacagcccg uacacaccgu gugcugggac accccacagu 240  
cagccgcaug gcucccccugu gccccagccc cuggcucccu cuguugaucc cggccccugc 300  
uccaggccuc acugugcaac ugcugcuguc acugcugcui cuggugccug uccaucccc 360  
gagguugccc cggaugcagg aggauucccc cuugggagga ggcucuucug ggaaagauga 420  
cccacugggc gaggaggauc ugcccaguga agaggauuca cccagagagg 470

<210> 56  
<211> 292  
<212> DNA  
<213> HUMAN

<400> 56  
gttttttta gacggagtct tgcatctgtc atgcccaggc tggagtagca gtggtgccat 60  
ctcggctcac tgcaagctcc acctcccgag ttcacgcccattttcctgcct cagcctcccg 120  
agttagctggg actacaggcg cccgccacca tgcccgcta atttttgta tttttggtag 180  
agacggggtt tcaccgtgtt agccagaatg gtctcgatct cctgacttcg tgatccaccc 240

gcctcgccct cccaaagttc tgggattaca ggtgtgagcc accgcacctg gc

292

<210> 57

<211> 262

<212> DNA

<213> HUMAN

<400> 57

tttcttttt gagacagggt cttgctctgt cacccaggcc agagtgaat ggtacagtct 60  
cagctcactg cagcctcaac cgcctcggt caaacatca tcccatattca gcctccttag 120  
tagctggac tacaggcaca tgccattaca cctggcta at tttttgtat ttcttagtaga 180  
gacagggtt gccatgttg cccgggctgg tctcgaactc ctggactcaa gcaatccacc 240  
cacctcagcc tcccaaaatg ag 262

<210> 58

<211> 2501

<212> DNA

<213> HUMAN

<220>

<221> misc\_feature

<222> (1)..(2501)

<223> region 5' to transcription initiation site as determined by RNase protection assay (nucleotide 3507 of Figures 2A-2F and of SEQ ID NO: 5), corresponding to region of SEQ ID NO: 5 and Figures 2A-2F from nucleotide (7) to nucleotide (2507), in which region some regulatory elements are probably situated.

<220>

<221> unsure what base is at position 1968

<222> (1968)

<223> unsure of base at position 1968, which is the same unknown base as that at position 1974 of SEQ ID NO. 5, i.e., the full-length MN genomic sequence, and of that unknown at position 1968 of SEQ ID NO: 90, and unknown at position 647 of SEQ ID NO: 110. That unknown base is in the 5' region flanking the transcription initiation site (3507) as determined by RNase protection assay.

<400> 58

tgttgactcg tgaccttacc cccaaaccctg tgctctctga aacatgagct gtgtccactc 60  
agggttaaat ggattaaggc cggtgcaaga tgtgctttgt taaacagatg cttgaaggca 120  
gcatgctcgt taagagtcat caccaatccc taatctcaag taatcaggga cacaacact 180  
gcggaaggcc gcagggtcct ctgccttagga aaaccagaga cctttgtca cttgtttatc 240  
tgaccttccc tccactattg tccatgaccc tgccaaatcc ccctctgtga gaaacaccca 300

agaattatca ataaaaaaaaat aaattaaaaa aaaaaaataca aaaaaaaaaaaa aaaaaaaaaaaa 360  
aaaagactta cgaatagttt ttgataaaatg aatagctatt ggtaaagcca agtaaatgat 420  
catattcaaa accagacggc catcatcaca gctcaagtct acctgatttgc atctctttat 480  
cattgtcatt ctggatttc actagatttag tcatacatcct caaaattctc ccccaagttc 540  
taattacgtt ccaaaccattt aggggttaca tgaagcttga acctactacc ttctttgctt 600  
tttagccatg agttgttagga atgatgagtt tacacccctac atgctgggaa ttaatttaaa 660  
cttacctct aagtcaaggttt ggtgcctt ggcttatttt ttagctaat tttgttagtt 720  
atggatgcac tgtaatctt gctatgatag tttcctcca cactttgcca cttagggtag 780  
gttagtactc agtttcagt aattgcttac ctaagaccct aagccctatt tctcttgta 840  
tggccttat ctgtaatatg ggcattatata atacaatata atttttggag tttttttgtt 900  
tgggttttgg tttgttttt tgagacggag tcttgcattt gtcatttttca ggctggagta 960  
gcagtggtgc catctcggtt cactgcaagc tccacccccc gagttcacgc cattttcctg 1020  
cctcagccctc ccgagtagct gggactacag gcgcggccca ccatgccccgg ctaattttt 1080  
gtatTTTgg tagagacggg gtttccacgt gttagccaga atggtctcga tctcctgact 1140  
tcgtgatcca cccgcctcggtt cctccaaag ttctgggattt acaggtgtga gccaccgcac 1200  
ctggccaattt tttgagtttctt tttaaagttaa aaatatgtct tgtaagctgg taactatgg 1260  
acatttcctt ttattaaatgt ggtgctgacg gtcattatagg ttctttttagg tttggcatgc 1320  
atatgctact tttgcagttt ctttcatcattt atttttctctt ctcatatttga agagcatgtt 1380  
atatcttttta gcttcacttg gcttaaaagg ttcttcattt agcctaacac agtgtcattt 1440  
ttggtaccac ttggatcata agtggaaaaa cagtcaagaa attgcacagt aataacttgg 1500  
tgtaagaggg atgattcagg tgaatctgac actaagaaac tcccctacccctt gaggtctgag 1560  
attcctctga cattgctgttataaggctttt tcctttgaca gcctgtgact gcggactattt 1620  
tttcttaagc aagatatgtt aaagttttgtt gagcctttt ccagagagag gtctcatatc 1680  
tgcatcaagt gagaacatat aatgtctgca tgtttccata tttcaggaat gtttgcttgt 1740  
gttttatgtt tttatataga cagggaaact tgttcctcag tgaccaaaaa gaggtgggaa 1800  
ttgttattttgg atatcatcat tggccacgc tttctgacccctt tgaaaacaat taagggttca 1860  
taatctcaat tctgtcagaa ttggatcataag aaatagctgc tatgtttctt gacattccac 1920  
ttggtaggaa ataaagaatgt gaaactcttc agttgggtgtg tgccctngt ttttttgc 1980  
tttccttctt actgtgtttaa aaaaaagtat gatcttgctc tgagaggtga ggcattctta 2040  
atcatgatct ttaaagatca ataataataat cctttcaagg attatgtctt tattataata 2100  
aagataattt gtcttaaca gaatcaataa tataatccct taaaggatta tatctttgtt 2160  
gggcgcagtg gctcacacccctt gtaatccctt cactttgggtt ggccaaagggtg gaaggatcaa 2220  
atttgcctac ttctatattttt tcttctaaag cagaattcat ctctcttccc tcaatatgtt 2280  
gatattgaca gggtttgcctt tcactcacta gattgtgagc tcctgctcag ggcaggtagc 2340  
gttttttgg tttgtttttt ttttctttt ttgagacagg gtcttgctt gtcacccagg 2400  
ccagagtgcata atggtagtactt ctcagctcac tgcagccctca accgcctcggtt ctcacccat 2460  
catccctt cagcctcctt agtagctggg actacaggca c 2501

<210> 59  
<211> 292  
<212> DNA  
<213> HUMAN

<220>  
<221> misc\_feature

<222> (1)

<400> 59

ttttttgag acggagtctt gcatctgtca tgcccaggct ggagtagcag tgggccatc 60  
tcggctcaact gcaagctcca cctcccgagt tcacgccatt ttcctgcctc agcctccga 120  
gtagctggga ctacaggcgc ccgcaccat gccggctaa tttttgtat ttttggtaga 180  
gacggggttt caccgtgtta gccagaatgg tctcgatctc ctgacttcgt gatccacccg 240  
cctcgccctc ccaaagttct gggattacag gtgtgagcca ccgcacctgg cc 292

<210> 60

<211> 262

<212> DNA

<213> HUMAN

<400> 60

ttctttttg agacagggtc ttgctctgtc acccaggcca gagtgcaatg gtacagtctc 60  
agctcaactgc agcctaacc gcctcggtc aaaccatcat cccatttcag cctcctgagt 120  
agctggact acaggcacat gccattacac ctggctaatt tttttgtatt tctagtagag 180  
acagggttg gccatgtgc ccgggctggc ctcgaactcc tggactcaag caatccaccc 240  
acctcagcct cccaaaatga gg 262

<210> 61

<211> 294

<212> DNA

<213> HUMAN

<400> 61

ttttttttg agacaaaactt tcactttgt tgcccaggct ggagtgcaat ggccgcgtct 60  
cggtcaactg caacctccac ctccccgggtt caagtgattc tcctgcctca gcctctagcc 120  
aagtagctgc gattacaggc atgcgccacc acgcccggct aattttgtt ttttttagtag 180  
agacggggtt tcgccatgtt ggtcaggctg gtctcgact cctgatctca ggtgatccaa 240  
ccaccctggc ctcccaaagt gctgggatta taggcgtgag ccacagcgcc tggc 294

<210> 62

<211> 276

<212> DNA

<213> HUMAN

<400> 62

tgcacgtctc tctgtcgccc aggctggagt gcagtgggt gatcttgggt cactgcaact 60  
tccgcctccc gggttcaagg gattctcctg cctcagcttc ctgagtagct ggggttacag 120  
gtgtgtgcca ccatgcccag ctaattttt tttgtatTTT tagtagacag gtttccacca 180  
tgttggtcag gtcggctca aactcctggc ctcaagtgtat ccgcctgact cagcctacca 240

aagtgctgat tacaagtgtg agccaccgtg cccagc

276

<210> 63

<211> 289

<212> DNA

<213> HUMAN

<400> 63

cgcggggcac ggtggctcac gcctgtaatc ccagcacttt gggaggccaa ggcaggtgga 60  
tcacgaggc aagagatcaa gaccatcctg gccaacatgg tgaaacccca tctctactaa 120  
aaatacgaaa aaatagccag gcgtggtggc gggtgcctgt aatcccagct actcgggagg 180  
ctgaggcagg agaatggcat gaacccggga ggcagaagtt gcagtgagcc gagatcgtgc 240  
cactgcactc cagcctgggc aacagagcga gactcttgtc taaaaaaaaa 289

<210> 64

<211> 298

<212> DNA

<213> HUMAN

<400> 64

aggctgggct ctgtggctta cgccataat cccaccacgt tgggaggctg aggtgggaga 60  
atggtttag cccaggagtt caagacaagg cggggcaaca tagtgtgacc ccatctctac 120  
aaaaaaaaacc ccaacaaaaac caaaaatagc cgggcatggt ggtatgcggc ctagtcccag 180  
ctactcaagg aggctgaggt gggaaagatcg cttgattcca ggagttttag actgcagtga 240  
gttatgatcc caccactgcc taccatcttt aggatacatt tatttattta taaaagaa 298

<210> 65

<211> 105

<212> DNA

<213> HUMAN

<400> 65

ttttttacat cttagtaga gacagggttt caccatattg gccaggctgc tctcaaactc 60  
ctgaccttgt gatccaccag cctcggcctc ccaaagtgtc gggat 105

<210> 66

<211> 83

<212> DNA

<213> HUMAN

<400> 66

cctcgaactc ctaggctcag gcaatccttt caccttagct tctcaaagca ctgggactgt 60

aggcatgagc cactgtgcct ggc

83

<210> 67  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 67  
agaaggtaag t

11

<210> 68  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 68  
Eggaggtgag a

11

<210> 69  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 69  
cagtcgtgag g

11

<210> 70  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 70  
ccgaggtgag c

11

<210> 71  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 71  
tggaggtacc a

11

<210> 72  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 72  
ggaaggtacag t 11

<210> 73  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 73  
agcaggtggg c 11

<210> 74  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 74  
cccaggtaca g 11

<210> 75  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 75  
tgctggtga g t 11

<210> 76  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 76  
atacaggggga t 11

<210> 77  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 77  
atacagggga t

11

<210> 78  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 78  
ccccaggcgca c

11

<210> 79  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 79  
acgcagtgc a

11

<210> 80  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 80  
tttcagatcc a

11

<210> 81  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 81  
ccccaggagg g

11

<210> 82

<211> 11  
<212> DNA  
<213> HUMAN

<400> 82  
tcacaggctc a

11

<210> 83  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 83  
cccttagctcc a

11

<210> 84  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 84  
ctccagtcga g

11

<210> 85  
<211> 12  
<212> DNA  
<213> HUMAN

<400> 85  
tcgcaggtga ca

12

<210> 86  
<211> 11  
<212> DNA  
<213> HUMAN

<400> 86  
acacagaagg g

11

<210> 87  
<211> 377

<212> PRT  
<213> HUMAN

<400> 87  
Gln Arg Leu Pro Arg Met Gln Glu Asp Ser Pro Leu Gly Gly Gly Ser  
1 5 10 15  
  
Ser Gly Glu Asp Asp Pro Leu Gly Glu Glu Asp Leu Pro Ser Glu Glu  
20 25 30  
  
Asp Ser Pro Arg Glu Glu Asp Pro Pro Gly Glu Glu Asp Leu Pro Gly  
35 40 45  
  
Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Glu Val Lys Pro Lys  
50 55 60  
  
Ser Glu Glu Glu Gly Ser Leu Lys Leu Glu Asp Leu Pro Thr Val Glu  
65 70 75 80  
  
Ala Pro Gly Asp Pro Gln Glu Pro Gln Asn Asn Ala His Arg Asp Lys  
85 90 95  
  
Glu Gly Asp Asp Gln Ser His Trp Arg Tyr Gly Gly Asp Pro Pro Trp  
100 105 110  
  
Pro Arg Val Ser Pro Ala Cys Ala Gly Arg Phe Gln Ser Pro Val Asp  
115 120 125  
  
Ile Arg Pro Gln Leu Ala Ala Phe Cys Pro Ala Leu Arg Pro Leu Glu  
130 135 140  
  
Leu Leu Gly Phe Gln Leu Pro Pro Leu Pro Glu Leu Arg Leu Arg Asn  
145 150 155 160  
  
Asn Gly His Ser Val Gln Leu Thr Leu Pro Pro Gly Leu Glu Met Ala  
165 170 175  
  
Leu Gly Pro Gly Arg Glu Tyr Arg Ala Leu Gln Leu His Leu His Trp  
180 185 190  
  
Gly Ala Ala Gly Arg Pro Gly Ser Glu His Thr Val Glu Gly His Arg  
195 200 205  
  
Phe Pro Ala Glu Ile His Val Val His Leu Ser Thr Ala Phe Ala Arg  
210 215 220

Val Asp Glu Ala Leu Gly Arg Pro Gly Gly Leu Ala Val Leu Ala Ala  
225 230 235 240

Phe Leu Glu Glu Gly Pro Glu Glu Asn Ser Ala Tyr Glu Gln Leu Leu  
245 250 255

Ser Arg Leu Glu Glu Ile Ala Glu Glu Gly Ser Glu Thr Gln Val Pro  
260 265 270

Gly Leu Asp Ile Ser Ala Leu Leu Pro Ser Asp Phe Ser Arg Tyr Phe  
275 280 285

Gln Tyr Glu Gly Ser Leu Thr Thr Pro Pro Cys Ala Gln Gly Val Ile  
290 295 300

Trp Thr Val Phe Asn Gln Thr Val Met Leu Ser Ala Lys Gln Leu His  
305 310 315 320

Thr Leu Ser Asp Thr Leu Trp Gly Pro Gly Asp Ser Arg Leu Gln Leu  
325 330 335

Asn Phe Arg Ala Thr Gln Pro Leu Asn Gly Arg Val Ile Glu Ala Ser  
340 345 350

Phe Pro Ala Gly Val Asp Ser Ser Pro Arg Ala Ala Glu Pro Val Gln  
355 360 365

Leu Asn Ser Cys Leu Ala Ala Gly Asp  
370 375

<210> 88  
<211> 34  
<212> DNA  
<213> HUMAN

<400> 88  
tagacagatc tacgatggct cccctgtgcc ccag 34

<210> 89  
<211> 34  
<212> DNA  
<213> HUMAN

<400> 89

attcctctag acagttaccg gctccccctc agat

34

<210> 90  
<211> 3532  
<212> DNA  
<213> HUMAN

<220>  
<221> misc\_feature which includes the MN gene promoter  
<222> (1)..(3532)  
<223> region including the transcription initiation site (nucleotide 3507  
of SEQ ID NO: 5 and of Figures 2A-2F) as determined by RNase protection  
assay, which region is inclusive of the MN gene promoter, and corresponds  
to nucleotide 7 to nucleotide 3538 of SEQ ID NO: 5 and of Figures 2A-2F.

220

<221> unsure what base is at position 1968

<222> (1968)

>223> unsure of the base at position 1968, which is the same unknown base at position 1974 of SEQ ID NO: 5 (the full-length MN genomic sequence), position 1968 of SEQ ID NO: 58 and position 647 of SEQ ID NO: 110. That unknown base is in the region that includes the transcription initiation site (nucleotide 3507 of SEQ ID NO: 5 and of Figures 2A-2F) as determined by RNase protection assay, which region is inclusive of the MN gene promoter.

<400> 90

tgttgactcg tgaccttacc cccaaccctg tgctctctga aacatgagct gtgtccactc 60  
agggttaaat ggattaaggg cggtgcaaga tgtgctttgt taaacagatg cttgaaggca 120  
gcatgcttgt taagagtcat caccaatccc taatctcaag taatcaggga cacaaacact 180  
gcggaaggcc gcagggtcct ctgccttagga aaaccagaga ccttggttca cttgtttatc 240  
tgaccttccc tccactattg tccatgaccc tgccaaatcc ccctctgtga gaaacaccca 300  
agaattatca ataaaaaaaaaa aaattttaaaa aaaaaataca aaaaaaaaaaaa aaaaaaaaaaa 360  
aaaagactta cgaatagttt ttgataaaatg aatagctatt ggttaagcca agtaaatgat 420  
catattcaaa accagacggc catcatcaca gctcaagtct acctgattt atctctttat 480  
cattgtcatt ctttggattc actagatttag tcacatcct caaaaattctc ccccaagttc 540  
taattacgtt ccaaacattt aggggttaca tgaagcttga acctactacc ttctttgtt 600  
ttgagccatg agttgttagga atgatgagtt tacaccttac atgctgggga ttaattttaaa 660  
ctttacctct aagtcagttg ggttagcctt ggcttatttt tgttagctaattttgttagtta 720  
atggatgcac tgtgaatctt gctatgatag ttttcctcca cactttgcca ctaggggttag 780  
gttaggtactc agtttcagt aattgcttac ctaagaccct aagccctatt tctcttgcac 840  
tggcctttat ctgtaatatg ggcataattta atacaatata atttttggag tttttttgtt 900  
tgtttgtttt tttgtttttt tgagacggag tcttgcattt gtcatttttttgcac 960  
gcagtgggtgc catctcggtc cactgcaagc tccacccccc gagttcacgc cattttcctg 1020  
cctcagccctc ccggacttagct gggactacag gcgcgcgcgc ccatgcggcgc ctaattttttt 1080

gtatTTTgg tagagacggg gttcaccgt gtagccaga atggtctcgta ttcctgtact 1140  
tcgtatcca cccgcctcggt cctccaaag ttctggatt acagggtgtga gccaccgcac 1200  
ctggccaatt tttgagtct tttaaagtaa aaatatgtct tgtaagctgg taactatgg 1260  
acatttcctt ttataatgt ggtgctgacg gtcataatagg ttcttttagg tttggcatgc 1320  
atatgctact tttgcagtc ctccattac attttctct cttcatttga agagcatgtt 1380  
atatcttta gcttcacttg gcttaaaagg ttctctcatt agcctaacad agtgtcattg 1440  
ttggtaccac ttggatcata agtggaaaaa cagtcaagaa attgcacagt aataactgtt 1500  
tgaagaggg atgattcagg tgaatctgac actaagaaac tcccctacct gaggtctgag 1560  
attcctctga cattgctgta tataggctt tccttgaca gcctgtgact gcggactatt 1620  
tttcttaagc aagatatgct aaagtttgt gagcctttt ccagagagag gtctcatatc 1680  
tgcataagt gagaacatata aatgtctgca tggttccata tttcaggaat gtttgctt 1740  
gttttatgct ttatataaga cagggaaact tggtcctcag tgacccaaaa gaggtggaa 1800  
ttgttattgg atatcatcat tggcccacgc tttctgaccc tggaaacaat taagggttca 1860  
taatctcaat tctgtcagaa ttggtacaag aaatagctgc tatgttctt gacattccac 1920  
ttggtaggaa ataagaatgt gaaactctc agttgggttg tgcctctngt tttttgcaa 1980  
tttccttctt actgtgttaa aaaaaagtat gatcttgctc tgagaggtga ggcattctta 2040  
atcatgatct taaagatca ataataataat ccttcaagg attatgtctt tattataata 2100  
aagataattt gtcttaaca gaatcaataa tataatccct taaaggatta tatcttgct 2160  
gggcgcagtg gctcacaccc gtaatcccag cactttgggt ggccaagggtg gaaggatcaa 2220  
atttgcttac ttcttatatta tcttctaaag cagaattcat ctctcttccc tcaatatgt 2280  
gatattgaca gggtttgcctc tcactcaacta gattgtgagc tcctgctcag ggcaggtagc 2340  
gtttttgtt ttgttttg ttttctttt ttgagacagg gtcttgctct gtcaccagg 2400  
ccagagtgca atggtacagt ctcagctcac tgagcctca accgcctcgg ctcaaaccat 2460  
catcccattt cagcctcctg agtagctggg actacaggca catgccatta cacctggcta 2520  
attttttgtt atttcttagta gagacagggt ttggccatgt tgcccggtt ggtctcgaac 2580  
tcctggactc aagcaatcca cccacctcag cctccaaaaa tgagggaccg tgtcttattc 2640  
atttccatgt cccttagtcca tagcccagtg ctggacctat ggtacta aataaatatt 2700  
tgttgaatgc aatagtaaat agcatttcag ggagcaagaa ctagattaac aaaggtggta 2760  
aaaggtttgg agaaaaaaaaat aatagttaa ttggctaga gtatgaggga gagtagtagg 2820  
agacaagatg gaaaggctc ttgggcaagg tttgaagga agttggaagt cagaagtaca 2880  
caatgtgcat atcggtgcag gcagtgggaa gccaatgaag gctttgagc aggagagtaa 2940  
tgtgttggaa aataaatata ggttaaacct atcagagccc ctctgacaca tacacttgct 3000  
tttcatcaat gctcaagttt gtctccacata taccattac ttaactcacc ctcggctcc 3060  
cctagcagcc tgccctaccc cttaacctgc ttccctgggtt agtcagggat gtatacatga 3120  
gctgcttcc ctctcagcca gaggacatgg gggccccag ctccctgcc ttcccttcc 3180  
tgtgcctgga gctggaaagc aggccagggt tagctgaggc tggctggcaa gcagctgggt 3240  
ggtgccaggagc agagcctgca tagtgccagg tggtgccctt ggttccaagc tagtccatgg 3300  
ccccgataac cttctgcctg tgcacacacc tgccctcac tccacccca tccttagctt 3360  
ggtatgggg agagggcaca gggccagaca aacctgtgag actttggctc catctctgca 3420  
aaagggcgct ctgtgagtca gcctgctccc ctccaggctt gctccctcccc caccagctc 3480  
tcgttccaa tgcacgtaca gcccgtacac accgtgtgct gggacacccc ac 3532

<210> 91  
<211> 204  
<212> DNA

<213> HUMAN

<400> 91  
cctgccccctc actccacccc catcctagct ttggtatggg ggagagggca cagggccaga 60  
caaacctgtg agactttggc tccatctctg caaaaggcg ctctgtgagt cagcctgctc 120  
ccctccaggc ttgctcctcc cccacccagc tctcgttcc aatgcacgta cagcccgtac 180  
acaccgtgtg ctgggacacc ccac 204

<210> 92

<211> 132

<212> DNA

<213> HUMAN

<400> 92

ggatcctgtt gactcgtgac cttaccccca accctgtgct ctctgaaaca tgagctgtgt 60  
ccactcaggg ttaaatggat taagggcggt gcaagatgtg ctttgttaaa cagatgcttg 120  
aaggcagcat gc 132

<210> 93

<211> 275

<212> DNA

<213> HUMAN

<400> 93

gcatagtgcc aggtggtgcc ttgggttcca agctagtcca tggcccgat aacttctgc 60  
ctgtgcacac acctgccccct cactccaccc ccatccttagc tttggtatgg gggagagggc 120  
acagggccag acaaacctgt gagactttgg ctccatctct gcaaaaggcg gctctgtgag 180  
tcagcctgct cccctccagg cttgctcctc ccccacccag ctctcgttc caatgcacgt 240  
acagcccgta cacaccgtgt gctgggacac cccac 275

<210> 94

<211> 89

<212> DNA

<213> HUMAN

<400> 94

ctgctccctt ccaggcttgc tcctccccca cccagctctc gtttccaatg cacgtacagc 60  
ccgtacacac cgtgtgctgg gacacccca 89

<210> 95

<211> 61

<212> DNA

<213> HUMAN

<400> 95

cacccagctc tcgtttccaa tgcacgtaca gcccgtacac accgtgtgct gggacacccc 60  
a 61

<210> 96

<211> 116

<212> DNA

<213> HUMAN

<400> 96

acctgccccct cactccaccc ccatccttagc tttggtatgg gggagagggc acagggccag 60  
acaaaacctgt gagactttgg ctccatctct gcaaaaagggc gctctgtgag tcagcc 116

<210> 97

<211> 36

<212> PRT

<213> HUMAN

<400> 97

Gly Glu Glu Asp Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp  
1 5 10 15

Pro Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro Gly Glu  
20 25 30

Glu Asp Leu Pro  
35

<210> 98

<211> 6

<212> PRT

<213> HUMAN

<400> 98

Gly Glu Glu Asp Leu Pro  
1 5

<210> 99

<211> 4

<212> PRT

<213> HUMAN

<400> 99  
Glu Glu Asp Leu  
1

<210> 100  
<211> 5  
<212> PRT  
<213> HUMAN

<400> 100  
Glu Glu Asp Leu Pro  
1 5

<210> 101  
<211> 6  
<212> PRT  
<213> HUMAN

<400> 101  
Glu Asp Leu Pro Ser Glu  
1 5

<210> 102  
<211> 7  
<212> PRT  
<213> HUMAN

<400> 102  
Glu Glu Asp Leu Pro Ser Glu  
1 5

<210> 103  
<211> 6  
<212> PRT  
<213> HUMAN

<400> 103  
Asp Leu Pro Gly Glu Glu  
1 5

<210> 104  
<211> 22  
<212> PRT  
<213> HUMAN

<400> 104  
Gly Gly Ser Ser Gly Glu Asp Asp Pro Leu Gly Glu Glu Asp Leu Pro  
1 5 10 15  
  
Ser Glu Glu Asp Ser Pro  
20

<210> 105  
<211> 25  
<212> PRT  
<213> HUMAN

<400> 105  
Gly Glu Glu Asp Leu Pro Ser Glu Glu Asp Ser Pro Arg Glu Glu Asp  
1 5 10 15  
  
Pro Pro Gly Glu Glu Asp Leu Pro Gly  
20 25

<210> 106  
<211> 24  
<212> PRT  
<213> HUMAN

<400> 106  
Glu Asp Pro Pro Gly Glu Glu Asp Leu Pro Gly Glu Glu Asp Leu Pro  
1 5 10 15  
  
Gly Glu Glu Asp Leu Pro Glu Val  
20

<210> 107  
<211> 7  
<212> PRT  
<213> HUMAN

<400> 107

Gly Glu Thr Arg Ala Pro Leu  
1 5

<210> 108  
<211> 7  
<212> PRT  
<213> HUMAN

<400> 108  
Gly Glu Thr Arg Glu Pro Leu  
1 5

<210> 109  
<211> 7  
<212> PRT  
<213> HUMAN  
  
<400> 109  
Gly Gln Thr Arg Ser Pro Leu  
1 5

<210> 110  
<211> 1247  
<212> DNA  
<213> HUMAN

<220>  
<221> misc\_feature  
<222> (1)..(1247)  
<223> region 5' to the transcription initiation site as determined by RNase protection assay (nucleotide 3507 of SEQ ID NO: 5 and of Figures 2A-2F) in which an activating element is localized, which region corresponds to nucleotide 1328 to nucleotide 2574 of SEQ ID NO: 5 and of Figures 2A-2F.

<220>  
<221> unsure what base is at position 647  
<222> (647)  
<223> unsure of the base at position 647, which is the same unknown base as that at position 1974 of SEQ ID NO: 5, and as that at position 1968 of SEQ ID NOS: 58 and 90. That unknown base at position 647 is in a region in which an activating element is localized and is 5' to the transcription initiation site.

<400> 110  
tatgctactt tttgcagtcc tttcattaca ttttctc ttcatttcaa gagcatgtta 60  
tatcttttag cttcaactgg cttaaaaggt tctctcatta gcctaaca ggttcattgt 120  
tggtaccact tggatcataa gtggaaaaac agtcaagaaa ttgcacagta atacttgtt 180  
gtaagaggga tgattcaggtaaat gatctgaca ctaagaaact cccctacctg aggtctgaga 240  
ttcctctgac attgctgtat ataggcttt ccttgacag cctgtgactg cgactattt 300  
ttcttaagca agatatgcta aagtttgtg agccttttc cagagagagg tctcataatct 360  
gcatcaagtg agaacatata atgtctgcat gttccatat ttcaggaatg tttgcttgc 420  
tttatgct ttatataagac agggaaaactt gttcctcagt gacccaaaag aggtggaaat 480  
tgttattgga tatcatcatt ggcccacgct ttctgacctt ggaaacaatt aagggttcat 540  
aatctcaatt ctgtcagaat tggtacaaga aatagctgt atgttcttgc acattccact 600  
tggttaggaaa taagaatgtg aaactcttca gttgggtgt gtccctngtt ttttgcatt 660  
ttccttccta ctgtgttaaa aaaaagtatg atcttgctct gagaggtgag gcattctaa 720  
tcatgatctt taaagatcaa taatataatc cttcaagga ttatgtctt attataataa 780  
agataatgg tctttaacag aatcaataat ataatccctt aaaggattat atcttgctg 840  
ggcgcgagtgg ctcacacctg taatcccagc actttgggtg gccaaagggtgg aaggatcaaa 900  
tttgcctact tctatattat cttctaaagc agaattcattc tctcttcctt caatatgtg 960  
atattgacag ggttgcctt cactcaactg attgtgagct cctgctcagg gcaggttagcg 1020  
tttttgcattttt ttgtttttgt tttctttt tgagacaggg tcttgctctg tcacccaggc 1080  
cagagtgc当地 tggtagcact tcagctcact gcagcctcaa ccgcctcggc tcaaaccatc 1140  
atccccatttc agccctctga gtagctggga ctacaggcac atgccattac acctggctaa 1200  
ttttttgtta tttcttagtag agacagggtt tggccatgtt gcccggg 1247

<210> 111  
<211> 17  
<212> DNA  
<213> HUMAN

<400> 111  
ctctgtgagt cagcctg 17

<210> 112  
<211> 23  
<212> DNA  
<213> HUMAN

<400> 112  
aggcttgctc ctccccccacc cag 23

<210> 113  
<211> 18  
<212> DNA  
<213> HUMAN

<400> 113  
agactttggc tccatctc 18

<210> 114  
<211> 20  
<212> DNA  
<213> HUMAN

<400> 114  
cactccaccc ccatccttagc 20

<210> 115  
<211> 26  
<212> DNA  
<213> HUMAN

<400> 115  
gggagagggc acagggccag acaaac 26

<210> 116  
<211> 20  
<212> PRT  
<213> HUMAN

<400> 116  
Gly Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly  
1 5 10 15

Gly Gly Gly Ser  
20